METADATA: Household Survey Main File

Dataset: household_survey_main_2024.csv

OVERVIEW

- **Purpose**: Main household-level data file for SADC household survey
- **Collection Period**: January - July 2024
- **Sampling Design**: Two-stage stratified cluster sample
- **Primary Sampling Units**: Enumeration Areas (EAs)
- **Secondary Sampling Units**: Households within EAs
- **Target Population**: All households in participating SADC countries
- **Sample Size**: Approximately 5,000 households

SAMPLING METHODOLOGY

- **Stage 1**: Selection of 250 EAs using PPS (probability proportional to size)
- **Stage 2**: Systematic selection of 20 households within each selected EA
- **Stratification**: By country and urban/rural classification
- **Weighting**: Design weights adjusted for non-response

VARIABLE DEFINITIONS

IDENTIFICATION VARIABLES

- **household_id** [String]: Unique household identifier (format: HH#####_###)
- **ea_id** [String]: Enumeration area identifier linking to EA master file
- **country** [String]: SADC member country name
- **stratum** [String]: Stratification variable (country_urban/rural)
- **urban_rural** [String]: Urban or Rural classification
- **province_code** [String]: Province identifier (P1-P8)
- **district_code** [String]: District identifier (D1-D25)

SURVEY ADMINISTRATION

- **interview_date** [Date]: Date of household interview (YYYY-MM-DD)
- **interview_mode** [String]: Data collection mode

- Face-to-face: In-person interview

- Telephone: Phone interview

- Web: Online self-administered

- **interview_duration_min** [Integer]: Interview length in minutes

- **interview_language** [String]: Language used for interview
  - English / Portuguese / French / Local

- **interview_result** [String]: Final interview outcome
  - Complete: All sections completed

  - Partial: Some sections missing

  - Refused: Household refused participation

  - Not Home: No contact after multiple attempts

HOUSEHOLD CHARACTERISTICS

- **household_size** [Integer]: Number of usual household members (1-15)

- **dwelling_type** [String]: Type of dwelling unit
  - House: Formal house/brick structure

  - Flat: Apartment/flat

  - Shack: Informal dwelling/shack

  - Traditional: Traditional dwelling/hut

  - Other: Other dwelling type

- **tenure_status** [String]: Ownership status of dwelling
  - Owned: Owned by household member

  - Rented: Renting from private/public

  - Free: Occupied free of charge

  - Other: Other arrangement

DWELLING MATERIALS

- **wall_material** [String]: Main material of external walls
  - Brick: Brick/concrete blocks

  - Mud: Mud/clay

  - Wood: Wood/planks

  - Iron: Corrugated iron/zinc

- Other: Other materials

- **roof_material** [String]: Main material of roof
  - Tile: Tiles
  - Iron: Corrugated iron/zinc sheets
  - Thatch: Thatch/grass
  - Concrete: Concrete/cement
  - Other: Other materials

SERVICES AND UTILITIES

- **water_source** [String]: Main source of drinking water
  - Piped: Piped water in dwelling/yard
  - Borehole: Borehole/protected well
  - Well: Unprotected well
  - River: River/stream/dam
  - Other: Other sources

- **electricity** [String]: Main source of electricity
  - Grid: Connected to national grid
  - Solar: Solar panels
  - Generator: Generator
  - None: No electricity

- **toilet_type** [String]: Type of toilet facility
  - Flush: Flush toilet (sewerage/septic)
  - Pit: Pit latrine
  - VIP: Ventilated improved pit
  - None: No facility/bush
  - Other: Other types

- **cooking_fuel** [String]: Main cooking fuel
  - Electricity / Gas / Paraffin / Wood / Charcoal / Other

DWELLING SIZE

- **n_rooms** [Integer]: Total number of rooms (excluding bathroom/toilet)
- **n_bedrooms** [Integer]: Number of rooms used for sleeping

HOUSEHOLD ASSETS (Binary 0/1)

- **has_radio**: Ownership of radio
- **has_tv**: Ownership of television
- **has_fridge**: Ownership of refrigerator
- **has_bicycle**: Ownership of bicycle
- **has_motorcycle**: Ownership of motorcycle
- **has_car**: Ownership of car/truck
- **has_computer**: Ownership of computer/laptop
- **has_internet**: Access to internet at home

ECONOMIC INDICATORS

- **monthly_income** [Numeric]: Total household monthly income (local currency)
- **monthly_expenditure** [Numeric]: Total household monthly expenditure
- **receives_grants** [Binary]: Household receives social grants (0=No, 1=Yes)
- **agricultural_hh** [Binary]: Engaged in agricultural activities
- **livestock_owned** [Binary]: Owns any livestock

RESPONDENT INFORMATION

- **respondent_age** [Integer]: Age of main respondent (18-85 years)
- **respondent_gender** [String]: Gender of main respondent (Male/Female)

SURVEY WEIGHTS

- **psu_weight** [Numeric]: Primary sampling unit (EA) weight
- **hh_weight** [Numeric]: Household-level weight
- **final_weight** [Numeric]: Final calibrated weight for analysis

DATA QUALITY

- **quality_check** [Binary]: Passed quality control (0=No, 1=Yes)
- **data_entry_date** [Date]: Date data entered into system

MISSING DATA CODES

- **NA**: Not available/Not applicable
- **-99**: Refused to answer

- **-88**: Don't know
- **Blank**: Skip pattern/Not asked

DATA QUALITY NOTES

1. Income and expenditure variables have approximately 5% item non-response
2. GPS coordinates missing for 2% of households due to equipment failure
3. Weights calibrated to 2022 census projections
4. Urban areas slightly oversampled to ensure adequate sample sizes
5. Panel households from 2023 survey included where possible

LINKAGE VARIABLES

- Links to **enumeration_areas_master.csv** via ea_id
- Links to **household_roster_2024.csv** via household_id
- Links to **panel_rotation_cohort_2023.csv** for panel households

USAGE NOTES

1. Always use survey weights for population-level estimates
2. Account for complex survey design in variance estimation
3. Check interview_result before including in analysis
4. Consider mode effects when interview_mode varies
5. Monetary values in local currency - apply exchange rates for comparison

VERSION INFORMATION

- **Version**: 1.0
- **Last Updated**: 2024-08-01
- **Contact**: sadc.survey@statistics.org
- **License**: Creative Commons Attribution 4.0

CITATION

SADC Statistical Unit (2024). Household Survey Main Dataset 2024.
Southern African Development Community, Gaborone, Botswana.